

CLUF: a Neural Model for Second Language Acquisition Modeling

Shuyao Xu, Jin Chen, Long Qin

Singsound Inc.

Introduction

Education systems that can adapt to the presenting of educational materials according to students' personal learning needs have great potential. Specifically, in the area of second language learning, Second Language Acquisition Modeling (SLAM) is the task to predict whether a second language learner would respond correctly in future exercises based on their learning history. In this work, we propose a neural network based system to utilize rich contextual, linguistic and user information. Our neural model consists of a Context encoder, a Linguistic feature encoder, a User information encoder and a Format information encoder (CLUF).

Method

We used in total four encoders to model the students' learning behavior:

- The context encoder consists of a character level LSTM encoder and a word level LSTM encoder.
- The linguistic feature encoder is also a LSTM model, where POS and syntactic dependency embedding are concatenated together and then fed into a multilayer LSTM unit.
- User encoder is fully-connected neural networks, which takes account of user id, users' nationality and other user related information.
- Format encoder is also fully-connected neural networks that encodes exercise format, session type, client type and time cost.

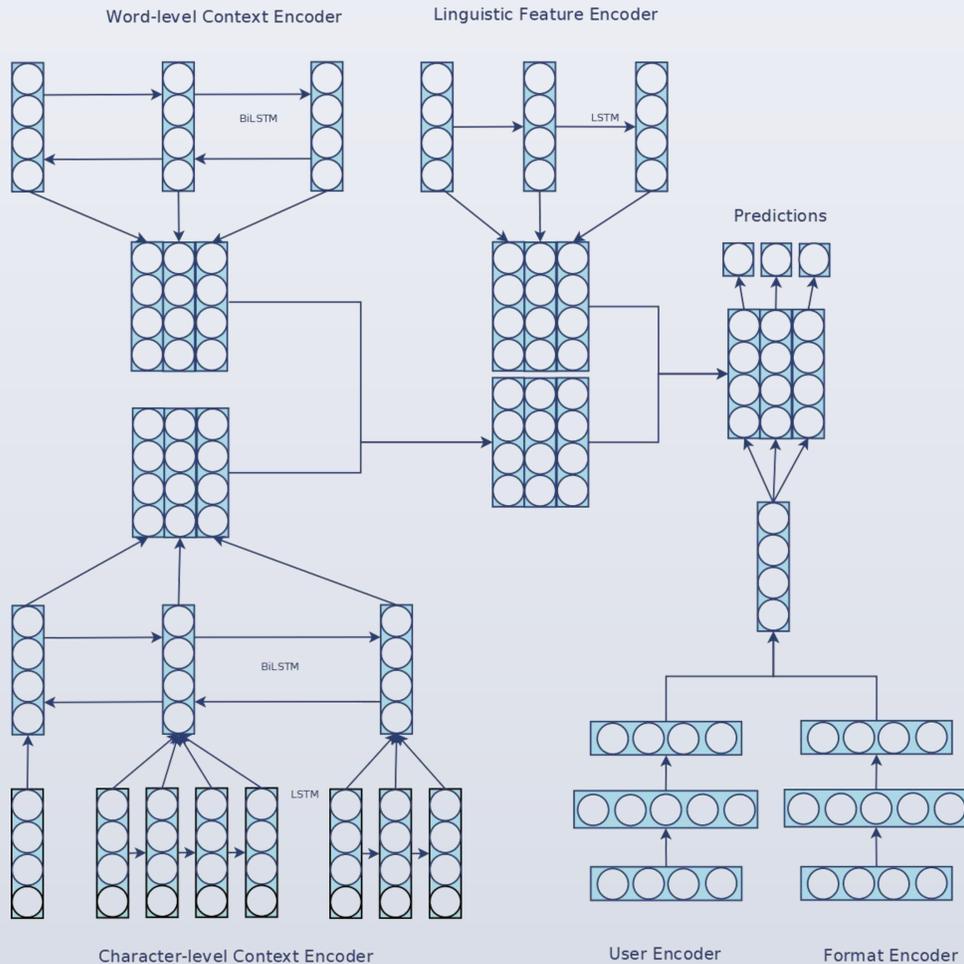


Figure 1. CLUF Architecture

The model is trained to minimize the following loss function:

$$Loss = -\frac{1}{N} \sum_{t=1}^N (\alpha y_t \cdot \log(p_t) + (1 - \alpha)(1 - y_t) \cdot \log(1 - p_t))$$

where α is the hyper parameter to balance the negative and positive samples and y_t is the label of the time step t and p_t is the prediction.

Category	Features
Context	Word Surface Form
Linguistic	Part of Speech, Syntactic Dependency Labels
User	User ID, Countries, Days in Course
Format	Client, Session Type, Exercise Format, Time Cost

Table 1. Feature Groups

Results

CLUF achieved AUC scores of 0.861, 0.835, 0.854, and F1 scores of 0.559, 0.524, 0.569 on the English track, Spanish track and French track respectively. As showed in Figure 2 and Table 2, CLUF gained significant improvement over the baseline model across the three tracks.

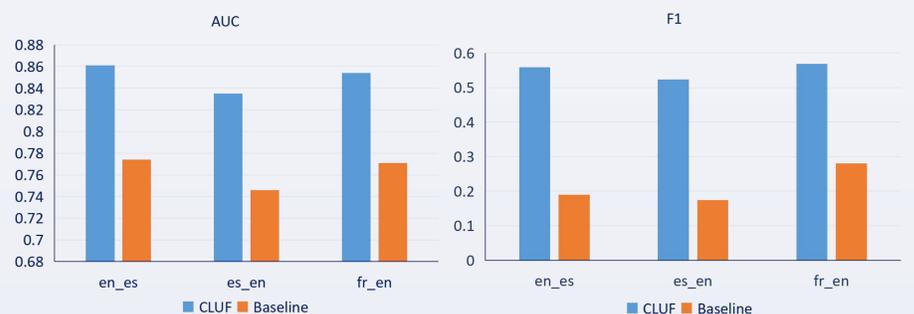


Figure 2. AUC and F1 Scores

Term	en_es	es_en	fr_en
Relative Impr (%)	11.24	11.93	9.72

Table 2. Relative AUC Improvement over Baseline

Discussion

The intuition behind CLUF is to factorize raw features into four independent parts: 1) word surface form models the word formation rules; 2) the linguistic encoder is to provide linguistic and syntactic dependency information; 3) the user part explores students' second language acquisition skills over time; 4) the format encoder measures the difficulty level of different exercises on various clients.

Model	AUC	F1
CLUF	0.846	0.554
LUF	0.775	0.446
CUF	0.843	0.552
CLF	0.813	0.501
CLU	0.779	0.467

Table 3. Results of Ablation Experiments.

Table 3 showed the performance of CLUF when excluding one of the context, linguistic, user and format encoder (LUF has no Context Encoder; CUF has no Linguistic encoder; CLF has no User Encoder; CLU has no Format Encoder). We can see that the performance drops substantially if we don't use the contextual or format features. On the other hand, excluding the linguistic features does not affect the performance much.

Conclusion

In this work, we proposed a neural based model, CLUF, for the SLAM task. We partitioned the features into four groups (Word surface features, Linguistic features, User features and Format features) which were encoded separately by different encoders. Experiments showed that our approach was effective and achieved very promising results. We further explored the contribution of each encoder and found that the context encoder was the most effective one, while the linguistic encoder was the least.